Factors influencing the discrimination and classification of prostate cancer cell lines by FTIR microspectroscopy†

T. J. Harvey, a E. Gazi, b A. Henderson, a R. D. Snook, a N. W. Clarke, b M. Brown b and P. Gardner a,b

Received 16th February 2009, Accepted 7th April 2009
First published as an Advance Article on the web 17th April 2009
DOI: 10.1039/b903249e

In this study we obtained Fourier transform infrared (FTIR) spectra of fixed prostate cell lines of differing types as well as the primary epithelial cells from benign prostatic hyperplasia (BPH). Results showed that by using multivariate chemometric analysis it was possible to discriminate and classify these cell lines, which gave rise to sensitivity and specificity values of >94% and >98%, respectively. Following on from these results the possible influences of different factors on the discrimination and classification of the prostate cell lines were examined. Firstly, the effect of using different growth media during cell culturing was investigated, with results indicating that this did not influence chemometric discrimination. Secondly, differences in the nucleus-to-cytoplasm (N/C) ratio were examined, and it was concluded that this factor was not the main reason for the discrimination and classification of the prostate cancer (CaP) cell lines. In conclusion, given the fact that neither growth media nor N/C ratio could totally explain the classification it is likely that actual biochemical differences between the cell lines is the major contributing factor.

1. Introduction

Prostate cancer (CaP) is the most commonly diagnosed cancer in males, with a mortality rate of about 10 000 people per year in the UK. The conventional diagnostic method for CaP diagnosis is the prostate specific antigen (PSA) test. PSA is a glycoprotein that is secreted by the secretory cells of the prostate gland, and is detected in the blood serum of patients.1 Elevated levels of PSA can be attributed to cancer; however, the threshold for a positive test is uncertain and a significant proportion of men will have cancer even with low PSA values. Thompson et al.2 found that out of 2950 men who never had PSA levels of more than 4 ng per milliliter, 449 patients (15.2%) were subsequently diagnosed with CaP. The major drawback of the PSA test is that it suffers from low specificities, resulting in a high false positive rate, leading to numerous unnecessary biopsies.3

Due to the drawbacks of current detection methods there is keen interest in the development of new minimally invasive screening tools for the identification of disseminated metastatic CaP cells in the peripheral blood or urinary systems.

Non-invasive, in-vitro FTIR spectroscopy is a powerful bio-analytical tool for the identification, classification and diagnosis of cancer cells or tissues.4–5 The transition of a normal cell to a diseased cell is accompanied by a change in a variety of biomolecules, that can be simultaneously and indiscriminately probed by FTIR microspectroscopy, yielding spectral signatures that enable the differentiation between normal and cancerous cells and tissues to be made.6–9 The use of FTIR spectroscopy with multivariate chemometrics to differentiate between different cell types has also been achieved. Such studies have included distinguishing between chemo-sensitive and chemo-resistant human chronic myelogenous leukaemia K562 cells by Gaigneaux et al.,10 between ectocervical and endocervical cells in cervical smears by Romeo et al.,11 and of particular relevance to this study, cancerous and non-cancerous bladder cells found in urine samples12.

Previously, we have reported preliminary findings that demonstrate the feasibility for FTIR microspectroscopy to differentiate benign from malignant prostate epithelial cells, and have shown that FTIR-photoacoustic spectra (FTIR-PAS) of CaP cells can be separated using principal component analysis (PCA).14 This opens up the possibility of using infrared and Raman spectroscopy to analyse prostate cells in urine samples, after they have been shed into the urine as a result of a prostate massage.15–17 To this end we have used optical tweezer Raman spectroscopy to analyse both live and fixed prostate and bladder cells in aqueous media and have shown that urine exposure prior to fixation does not inhibit classification.18–21

In both our preliminary IR work and our later Raman work it has been assumed that cell classification is based on subtle differences in biochemistry associated with the cell phenotype. This assumption is generally held by practitioners in the field but has recently been called into question by Romeo et al.22,23 In an elegant study of human oral mucosa cells and canine cervical cells the authors showed that the IR spectra of human oral mucosa cells grouped with those of the canine cervical cells of estrus dogs. Furthermore, they were able to demonstrate that the principal factor influencing the separation of cells is the
nucleus-to-cytoplasm (N/C) ratio. This seminal study, which showed that the difference between human and canine cells was spectrally less discriminating than the N/C ratio, has potentially serious ramifications for the field.

In this present study we use infrared microspectroscopy to record infrared spectra of four closely related prostate cell types and classify them using established chemometric methods. We then examine two possible factors, other than the subtle biochemical differences associated with differing phenotype that could be responsible for observed discrimination and classification. These factors are the influence of (i) the culture media and (ii) the nucleus-to-cytoplasm ratio.

2. Experimental methods

This section is subdivided into two: experimental methods for classification and experimental methods for determining factors responsible for classification.

2.1 Classification

2.1.1 Preparation of cell lines and primary cells. For this current research two malignant (LNCaP and PC-3), a non-malignant primary cell line (BPH) and a transformed normal (PNT2-C2) cell type were studied. The lymph node and bone marrow metastasis, LNCaP and PC-3, respectively are two of the most studied CaP cell lines and are commercially available. LNCaP cells are well differentiated and androgen-sensitive whilst PC-3 cells are less well differentiated, are androgen-independent and display highly aggressive and invasive phenotypic behavior. PNT2-C2 is a normal prostate epithelial cell transfected with the SV40 virus expressing the large T antigen, resulting in cell immortality, and as such acts as a model ‘cancerous’ cell. BPH (benign prostatic hyperplasia) is the most common non-malignant disorder of the prostate and results in enlargement of the glands, mainly in the transition zone of the prostate gland.

Five slides of each cell line were prepared using MirrIR plates (Kevley Technologies). MirrIR slides are widely used within the biospectroscopy community since they are inert, non-cytotoxic and a fraction of the cost of IR-transparent substrates. LNCaP cells and PNT2-C2 cells were cultured in RPMI 1640 with 10% v/v foetal calf serum (FCS) and 2 mM L-glutamine. PC-3 cells were cultured in Ham’s F12 with 7% v/v FCS and 2 mM L-glutamine.

2.1.2 FTIR microspectroscopy. A minimum of 50 FTIR spectra per cell line for each batch were collected in reflection mode using a Nicolet FTIR spectrometer coupled to a Nic-Plan microscope and equipped with an MCT liquid N2-cooled detector. The IR passes through the cell twice, so this mode of collection is often termed transflection. Each spectrum was acquired with a sampling aperture of 60 μm × 60 μm and 256 co-added scans at a resolution of 4 cm⁻¹. The spectra were ratioed against a background spectrum taken on a clean MirrIR slide using identical parameters as described for the sample spectrum.

2.1.3 Data processing. All spectra were converted into the 1st derivative using OMNIC software. A separate copy of the data was converted into the 2nd derivative. The spectral region 800–1481 cm⁻¹ was chosen for analysis. All other parts of the spectrum were discarded, since the aforementioned spectral range contains the so-called fingerprint region, which proved to be the most diagnostically significant. The data were then imported into MATLAB (Mathworks Inc.) and subjected to vector normalisation, mean-centring and PCA. Vector normalisation is used to account for different amounts of material being sampled and is a standard preprocessing method employed by many leading groups in the field. Principal components (PCs) generated from PCA were then imported into SPSS software (SPSS Inc.) for linear discriminant analysis (LDA), subsequently referred to as PC-LDA.

2.2 Factors responsible for classification

2.2.1 Effect of growth media. PC-3 and LNCaP cells were grown onto MirrIR slides in both Ham’s F12 and RPMI 1640 media respectively using the methods outlined in the preceding text. FTIR spectra were recorded using the same instrumentation setup and methods as previously reported. Resultant spectra were subjected to calculation of the 1st derivative, followed by vector normalisation, mean-centring and PCA.

2.2.2 N/C ratio measurements

2.2.2.1 Fluorescence staining of cells. The cytoplasm and nucleus of prostate cells were differentially stained with fluorescing biomarkers, and viewed under an Olympus BX51 fluorescence microscope. Cellular membranes, especially the ER and Golgi, were stained with BODIPY TR Ceramide to define the cellular cytoplasm (Invitrogen), whilst the nucleus was stained by DAPI (Vectorshield).

Formalin-fixed cells on MirrIR slides, which were previously used to collect the FTIR spectra were re-hydrated in HBSS (Hank’s Balanced Salt Solution) with 1% HEPES (salt balanced buffer solution) for 20 minutes. The slides were then incubated at 4°C for 35 minutes. Following this, excess BODIPY solution was siphoned off and 1–2 drops of DAPI stain was added to the slides. A cover slip was placed on the slide and sealed around the edges with nail-varnish. The slides

The cells were briefly washed in distilled water to remove residual PBS from their surface, dried under ambient conditions and stored in a desiccator until FTIR analysis. To ascertain experimental reproducibility two further cell batches were prepared on different days using the same method as described above.

2.1.2 FTIR microspectroscopy. A minimum of 50 FTIR spectra per cell line for each batch were collected in reflection mode using a Nicolet FTIR spectrometer coupled to a Nic-Plan microscope and equipped with an MCT liquid N2-cooled detector. The IR passes through the cell twice, so this mode of collection is often termed transflection. Each spectrum was acquired with a sampling aperture of 60 μm × 60 μm and 256 co-added scans at a resolution of 4 cm⁻¹. The spectra were ratioed against a background spectrum taken on a clean MirrIR slide using identical parameters as described for the sample spectrum.

2.1.3 Data processing. All spectra were converted into the 1st derivative using OMNIC software. A separate copy of the data was converted into the 2nd derivative. The spectral region 800–1481 cm⁻¹ was chosen for analysis. All other parts of the spectrum were discarded, since the aforementioned spectral range contains the so-called fingerprint region, which proved to be the most diagnostically significant. The data were then imported into MATLAB (Mathworks Inc.) and subjected to vector normalisation, mean-centring and PCA. Vector normalisation is used to account for different amounts of material being sampled and is a standard preprocessing method employed by many leading groups in the field. Principal components (PCs) generated from PCA were then imported into SPSS software (SPSS Inc.) for linear discriminant analysis (LDA), subsequently referred to as PC-LDA.

2.2 Factors responsible for classification

2.2.1 Effect of growth media. PC-3 and LNCaP cells were grown onto MirrIR slides in both Ham’s F12 and RPMI 1640 media respectively using the methods outlined in the preceding text. FTIR spectra were recorded using the same instrumentation setup and methods as previously reported. Resultant spectra were subjected to calculation of the 1st derivative, followed by vector normalisation, mean-centring and PCA.

2.2.2 N/C ratio measurements

2.2.2.1 Fluorescence staining of cells. The cytoplasm and nucleus of prostate cells were differentially stained with fluorescing biomarkers, and viewed under an Olympus BX51 fluorescence microscope. Cellular membranes, especially the ER and Golgi, were stained with BODIPY TR Ceramide to define the cellular cytoplasm (Invitrogen), whilst the nucleus was stained by DAPI (Vectorshield).

Formalin-fixed cells on MirrIR slides, which were previously used to collect the FTIR spectra were re-hydrated in HBSS (Hank’s Balanced Salt Solution) with 1% HEPES (salt balanced buffer solution) for 20 minutes. The slides were then incubated at 4°C for 35 minutes. Following this, excess BODIPY solution was siphoned off and 1–2 drops of DAPI stain was added to the slides. A cover slip was placed on the slide and sealed around the edges with nail-varnish. The slides
were then viewed using the Olympus BX51 fluorescence microscope and captured using an attached video/camera system. For each slide, two images were obtained, one showing the BODIPY-stained cytoplasm and one showing the DAPI-stained nucleus. The text in Section 3.3.2 outlines the process by which this data was used to obtain the N/C ratio results.

2.2.2.2 AFM measurements. Atomic force microscope (AFM) imaging was utilised to obtain a 3-D representation of the cells. Cells were cultured onto glass slides using the same method outlined in Section 2.1. A Multimode AFM (Digital Instruments) was used in conjunction with a Zeiss optical microscope and coupled to a Nanoscope IIIa AFM controller. To image the cells in contact mode a cantilever of length 200 μm with a silicon nitride probe tip of spring constant 0.06 N m⁻¹ and a tip height of 2.5–3.5 μm was used. Each image was raster scanned with an image field of between 48 and 52 μm at a scan rate of between 0.10 and 0.22 Hz and constructed from 512 × 512 pixels. Use of the AFM image data to obtain N/C results is discussed in Section 3.3.2.

3. Results and discussions

3.1 Analysis of raw, 1st derivative and 2nd derivative FTIR spectra of prostate cell lines

Fig. 1(a) shows the raw mean absorbance spectra for the 800–1800 cm⁻¹ spectral region. The mean represents all data collected from all three batches analysed (167 BPH, 172 LNCaP, 167 PC-3 and 170 PNT2-C2 spectra). Average standard deviation of normalised absorbance intensities for the 800–1800 cm⁻¹ region gave results (in increasing spectral variance): LNCaP (0.036) < BPH (0.039) < PNT2-C2 (0.046) < PC-3 (0.095). Fig. 1(b) and 1(c) show the mean 1st and 2nd derivative vector normalised spectra respectively for the 800–1481 cm⁻¹ region (the region used for subsequent PCA analysis). A similar spectral region was previously used to correlate FTIR spectra with the Gleason grade. The derivative plots highlight some differences between the spectra, with LNCaP displaying the most prominent differences, most notably in the 1050–1150 cm⁻¹ region.

3.2 Multivariate chemometric analysis of FTIR spectra obtained from prostate cells

3.2.1 PCA. All 1st and 2nd vector normalised derivative spectra were subjected to PCA. PCA is an unsupervised exploratory technique, which reduces the large number of variables inherent within a spectrum to a few principal components (PCs). In doing so it enables the relationship between spectra to be easily visualised on a 2-D or 3-D scatter plot. Fig. 2(a) and 2(b) show the PCA plots for both the 1st and 2nd derivative FTIR spectra (the plots contain the same number of spectra for each cell line as quoted above). Fig. 2(a) shows evidence of separation of the four cell lines into four defined quadrants, with separation occurring on both the PC1 and PC3 axes. The plot also shows that LNCaP and BPH appear to have the greatest separation of any cell line pair, whilst there is some overlap of the PC-3 and PNT2-C2 spectra. For the 2nd derivative plot (Fig. 2(b)) separation of cell lines occurs only on the PC2 axis, with LNCaP and BPH displaying the greatest separation. Of further interest is the fact that LNCaP and PNT2-C2 separate despite being grown in the same media. Therefore, this strongly suggests that using different growth media does not contribute to the discrimination. The role of growth media is investigated more comprehensively later in the paper.

3.2.2 Prostate cell line classification using PC-LDA modelling. A diagnostic classifier algorithm was constructed using LDA, which has previously been used to classify breast tumours and prostate tumour cells. This supervised method searches variables which contain the smallest intra-group separation and the largest inter-group separation and constructs a linear combination of variables to discriminate between the groups. However, one drawback of the LDA method is that the numbers of spectra in the smallest group have to be at least twice

![Fig. 1](image-url) Mean absorbance spectrum (a), 1st derivative spectrum (b) and 2nd derivative spectrum (c) for each of the four prostate cell lines.
scores are shown in Table 1, and show that 100% of blind LNCaP indicating that the training set is not over-fitted.) Classification blind-set clusters are of a similar size to those of the training set are incorrectly assigned to the PC-3 cluster. (Importantly all the (Fig. 3(b)) cluster mostly with the BPH spectra, and only a few generally correctly assigned. For example, the blind BPH spectra LDF2 separates LNCaP from BPH, PC-3 and PNT2-C2. PCA plots (Fig. 2). LDF1 separates BPH, PC-3 and PNT2-C2, whilst LDF2 separates LNCaP from BPH, PC-3 and PNT2-C2. Fig. 3(b–e) shows the same PC-LDA plots but with blind BPH, LNCaP, PC-3 and PNT2-C2 spectra respectively projected into the PC-LDA classifier (all spectra added one cell line at a time).

Observation of the plots shows that the blind spectra are generally correctly assigned. For example, the blind BPH spectra (Fig. 3(b)) cluster mostly with the BPH spectra, and only a few are incorrectly assigned to the PC-3 cluster. (Importantly all the blind-set clusters are of a similar size to those of the training set indicating that the training set is not over-fitted.) Classification scores are shown in Table 1, and show that 100% of blind LNCaP and PNT2-C2 spectra are correctly assigned. These results were used to calculate the sensitivity and specificity values, which are displayed in Table 2.

Sensitivities and specificities values are often quoted for clinical diagnostic models and give an indication of the diagnostic performance of the model.42 In the case of LNCaP, the sensitivity value represents the probability that the PC-LDA model will give a positive score among all LNCaP cell lines. The specificity value represents the probability that the PC-LDA model will be negative among all cell lines that are not LNCaP.

As a consequence of the large separation of the four clusters in the training set PC-LDA template, blind test spectra can be projected into the model with a high degree of accuracy as shown by the high sensitivity and specificity results. The highest sensitivity results were achieved for LNCaP and PNT2-C2 cells, while the highest specificity result was achieved for BPH. The high values for LNCaP and BPH can be attributed to the fact that they are most separated from any other cell line type, while PC-3 and PNT2-C2 cluster close together.

The high sensitivity and specificity values indicates that this PC-LDA could classify unknown CaP cells with a high degree of accuracy, and therefore with promising diagnostic potential. These results are comparable to those obtained by Crow et al., who used Raman microspectroscopy to classify CaP cell lines.41 In the work by Crow et al. similar high prediction values were obtained, but using internal cross-validation rather than blind testing.

3.3 Factors responsible for classification

In the preceding text we have reported that FTIR spectra obtained from prostate cells can be discriminated and classified with a high degree of accuracy using both PCA and PC-LDA. In the following section we investigate two possible factors responsible for these results: growth media and N/C ratio.

3.3.1 Growth media. The FTIR spectra recorded in this research were obtained from cells cultured in different growth media. LNCaP and PNT2-C2 cells were grown in RPMI 1640, PC-3 cells in Ham’s F12 and BPH cells in K-SFM. These are the growth media recommended by The European Collection of Cell Cultures (ECACC).43 It could be argued therefore that discrimination of the cell lines is not due to intrinsic cellular biochemistry, but rather due to media-induced changes, e.g. cells grown in a lipid-rich medium may exhibit a more lipid-rich biochemistry compared with similar cells grown in lipid-poor media. The aim of this study was to determine the effect of using the different growth media on the discrimination of prostate cells.

Fig. 4 displays the PCA plot of LNCaP and PC-3 grown in both RPMI 1640 and Ham’s F12. Results show two clear clusters, separated based on the cell type and not growth medium. These results show that for LNCaP and PC-3 cells the growth medium they are cultured in does not influence significantly the resulting spectra. If growth medium did significantly influence the spectra then it would have been expected that the spectra would cluster based on the medium. The intrinsic biochemistry of the cell is therefore not determined by the growth medium to the extent that it can be detected by FTIR spectroscopy.
To further confirm our findings above, one-way ANOVAs were carried out on the PC2 component scores from the PCA plot at the 95% confidence level. The following combinations were tested:

Fig. 3  PC-LDA training set template (a), BPH spectra projected blind (b), LNCaP spectra projected blind (c), PC-3 spectra projected blind (d) and PNT2-C2 spectra projected blind (e). These results are based on 1st derivative data.

| Table 1  PCA-LDA scores for blind testing of classifier model |
|---------|---|---|---|---|
| Desired class | Classification score (%) |
| BPH | LNCaP | PC-3 | PNT2-C2 |
| BPH | 94.4 | 0.9 | 4.7 | 0 |
| LNCaP | 0 | 100 | 0 | 0 |
| PC-3 | 0 | 0 | 99.1 | 0.9 |
| PNT2-C2 | 0 | 0 | 0 | 100 |

| Table 2  Sensitivity and specificity values for the PC-LDA classifier model |
|---------|---|---|---|---|
| Sensitivity (%) | BPH | LNCaP | PC-3 | PNT2-C2 |
| 94.4 | 100 | 99.1 | 100 |
| Specificity (%) | 100 | 99.7 | 98.4 | 99.7 |

To further confirm our findings above, one-way ANOVAs were carried out on the PC2 component scores from the PCA plot at the 95% confidence level. The following combinations were tested:
LNCaP/RPMI 1640 vs. LNCaP/Ham’s F12 (i.e. these spectra cluster in Fig. 4). The PC2 scores were calculated not to be significantly different ($p = 0.77$).

PC-3/RPMI 1640 vs. PC-3/Ham’s F12 (i.e. these spectra cluster in Fig. 4). The PC2 scores were calculated not to be significantly different ($p = 0.21$).

LNCaP/RPMI 1640 + LNCaP/Ham’s F12 vs. PC-3/RPMI 1640 + PC-3/Ham’s F12 (these spectra cluster based on cell types in Fig. 4). The PC2 scores were calculated to be significantly different ($p = 0.02$).

These results confirm observational findings that separation is based on cell type and not the media they were grown in.

### 3.3.2 N/C ratios.

The N/C ratio is a common parameter used by cell biologists to characterise changes in cells undergoing transformations and to determine cell health.\(^4\) Research by Romeo et al. suggested that the PCA separation between FTIR spectra obtained from human oral mucosa cells and canine cervical cells was a consequence of the N/C ratio.\(^2\) This conclusion was reached by incorporating spectra of human oral cells and canine cervical cells from both estrus and non-estrus dogs into the same PCA algorithm. The researchers found that the human oral cells and the cells from dogs in estrus clustered together and were well separated from cells of non-estrus dogs. The explanation given for this observation was that the immature cervical cells from non-estrus dogs are smaller compared with the cells of estrus dogs, and the cells of estrus dogs were roughly the size of the human oral cells. Because the amounts of nucleus and cytoplasm probed will vary with cell size it was postulated that the N/C ratio was responsible for the discrimination between these cell types.

In the light of the work by Romeo et al. it is important to establish if the N/C ratio is the dominating factor influencing the discrimination and classification of the prostate cell lines.

#### 2-D Calculations.

By the use of ImageJ software (downloaded from http://rsbweb.nih.gov/ij/ on 11th May 2006) the nucleus and cytoplasm area in an isolated cell was determined, together with the area of the nucleus only. For cell clusters, the drawing tool option in ImageJ enabled the total area (cytoplasm and nuclei) of the cluster to be acquired. The total number of cells measured per cell line was between 70 and 90. Following this, the cytoplasmic area alone was calculated by subtracting the total cell area per plate from the sum of areas of all nuclei per plate. The N/C ratio for each plate was then simply calculated by dividing nucleus area by cytoplasm area.

Fig. 5 shows a typical PC-3 image used to calculate the N/C ratio, as described above. The image consists of (a) BODIPY-stained membranes delineating the cytoplasm, (b) DAPI-stained nuclei and (c) a combined image showing stained cytoplasm and nuclei (colours were added retrospectively using Jasc Paint Shop Pro).

#### 3-D Calculations.

From the AFM image data the cell height and total cell volume were obtained using SPIP software (Image Metrology). The cell volume was obtained by drawing a box round the whole cell, with the software computing the volume. The mean heights and volumes from 5 cells per cell line were calculated. Fig. 6(a) shows an AFM image of a typical BPH cell used to obtain volume information for 3-D measurements. Fig. 6(b) shows the cross-section profile used to obtain height information.

From the fluorescence stained images of the nucleus a total nucleus area per slide was obtained (these are the same slides as referred to for the 2-D area calculations). By counting the number of cells on the slide an average nucleus area and thus the average nucleus radius, $R_n$, was obtained.

To a rough approximation it was assumed that the nucleus takes the shape of an ellipsoid, with volume given by eqn (1), where $a$, $b$ and $c$ are the axis radii:
In the case of eqn (1), \(a\) and \(b\) were considered to be \(R_n\) and \(c\) the cell height obtained from the AFM images. From the calculated nucleus volume per cell, together with the mean cell volume measurements from the AFM data, an N/C ratio value could be obtained for each slide. Table 3 shows the N/C ratio results for three different images (slides) per cell type for both 2-D (area) and 3-D (volume) measurements.

Results show that for both 2-D and 3-D measurements, BPH cells have the lowest N/C ratio while PNT2-C2 cells the highest. A large N/C ratio for SV40 virus transfected cells, such as PNT2-C2, may be typical, as Isom et al.\(^45\) found an increase in N/C ratio for transfected rat hepatocytes cells. It is apparent that N/C ratios from area measurements are significantly higher than those from volume measurements, but the overall patterns are the same. Results show clear differences in the N/C ratio between cell lines, which could, in theory, be used to explain the classifications observed using chemometric analysis (PCA and PC-LDA results).

Chemometric results indicate that LNCaP and BPH have the largest separation between any cell line type, and PNT2-C2 and PC-3 the smallest (refer back to Fig. 2 and 3). However, this contradicts with the N/C ratio results which show that BPH and PNT2-C2 have the largest separation. If the N/C ratio were the dominant factor responsible for the discrimination, a direct correlation between the N/C ratio and the chemometric results might have been expected. Because this correlation does not exist, it can be suggested that the N/C ratio is not the dominant factor and does not explain the discrimination.

To further validate this argument, simulated spectra consisting of varying amounts of nucleus and cytoplasm were generated. Highly resolved spectral maps of all four cell lines were obtained using a Perkin Elmer Spotlight FTIR spectrometer with a linear array detector. From these maps, spectra of the nucleus and cytoplasm were extracted, and an average nucleus and cytoplasm spectrum from each cell determined. Areas of the map showing high intensity of the phosphate moiety at 1240 cm\(^{-1}\) were chosen for extracting nucleus spectra. For extracting cytoplasm spectra, intensities of the ratio of integrated lipid stretch peaks (2850–3000 cm\(^{-1}\)) against integrated amide I and II peaks (1500–1700 cm\(^{-1}\)) were used.

The two spectra representing the nucleus and cytoplasm were then mixed in various proportions from 0% nucleus to 100% nucleus using OMNIC software. This method was carried out for two cell maps per cell line, resulting in eight spectral sets of varying N/C ratios. PCA was performed on the eight cells using simulated spectra representing 10 : 90, 40 : 60, 60 : 40 and 90 : 10 N/C ratios (Fig. 7). This plot shows that spectra from each of the four cell lines cluster together, with distinct separation between cell line types. There is no clustering of spectra based on their N/C ratio and the separation between cell line types is much greater than the separation between spectra of different N/C ratios from the same cell.

Whilst these results are interesting, simulating spectra in this way generates an imperfect model with two important limitations. Firstly, these simulated spectra are artificial and so do not represent a real situation of cells with different N/C ratios. Secondly, spectra obtained from the nucleic region are probably not ‘pure’ nucleus and might contain some cytoplasmic contributions.

### Table 3 2-D (area) and 3-D (volume) N/C ratio results

<table>
<thead>
<tr>
<th>Cell type</th>
<th>N/C ratio 2-D (area)</th>
<th>N/C ratio 3-D (volume)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slide 1</td>
<td>Slide 2</td>
</tr>
<tr>
<td>BPH</td>
<td>0.44</td>
<td>0.39</td>
</tr>
<tr>
<td>LNCaP</td>
<td>1.17</td>
<td>1.12</td>
</tr>
<tr>
<td>PC-3</td>
<td>0.86</td>
<td>0.66</td>
</tr>
<tr>
<td>PNT2-C2</td>
<td>2.40</td>
<td>1.61</td>
</tr>
</tbody>
</table>

\[ V = \frac{4}{3} \pi abc \]

Fig. 6 AFM image of a BPH cell (a), where the white line indicates the horizontal cross-section area in which cell height was obtained (b).
These findings, however, together with the results from the 2-D and 3-D calculations suggest that for these prostate cells the N/C ratio does not explain the results seen in the PCA and PC-LDA results. However, we cannot say whether these conclusions hold for all cell lines, which have been studied by other researchers.

4. Conclusions

In this paper the application of FTIR microspectroscopy to analyse and classify prostate cell lines has been demonstrated. FTIR spectra from 3 different batches of PC-3, LNCaP, PNT2-C2 and BPH were obtained in reflection mode and subjected to unsupervised and supervised multivariate chemometric analysis. It was shown that PCA alone could separate the four cell types using either 1st or 2nd derivative spectra. Optimal clustering was achieved when the PCs were further analysed by an LDA algorithm. Blind testing of the PC-LDA classifier achieved very promising prediction values (>94% and >98% for sensitivities and specificities, respectively). These results demonstrate that FTIR spectroscopy can be used to discriminate and classify prostate cell lines with a high degree of accuracy.

As well as carrying out chemometric analysis on the FTIR spectra possible factors for the discrimination and classifications observed in the PCA and PC-LDA plots were investigated. Firstly, it was ascertained that using different growth media in the cell culturing process did not significantly influence the results.

Secondly, the influence of the N/C ratio on classification was ascertained. Results show differences in the N/C ratio between cell lines, which could explain the classification. However, they did not correlate with the chemometric results. These results were confirmed by PCA plots of spectral simulations which show that spectra cluster by cell type and not by the N/C ratio. While not conclusive, these results would indicate that the N/C ratio does not overtly influence classification of prostate cells.

Whilst there is likely to be some correlation between all factors that may influence the spectra, i.e. growth media, N/C ratio and inherent biochemical differences, the results presented here strongly suggest that neither growth media nor N/C ratio play a dominant role. Could morphological differences influence classification? This is an important question that has not been addressed directly in this paper. It could be hypothesized that physical differences between the cells result in different degrees of scattering which could influence spectra discrimination. Although Mie scattering in particular undoubtedly occurs to some degree, the fact that the spectra represent an average over many cells seems to limit this effect on the spectra. In addition, Raman spectra of prostate cells in a liquid environment obtained in the optical tweezer system showed a similar degree of separation as the infrared spectra presented here. Since the scattering regimes for these two experimental setups are completely different (different sample environment, different wavelengths of probe beam) it is reasonable to assume that scattering is also not the dominant factor. It is most likely therefore that the inherent biochemical differences between the prostate cell lines is the single biggest factor that influences the classification results.

Acknowledgements

We acknowledge The Prostate Cancer Research Foundation for financing this project (T. H.). We thank Dr Steve Bagley (Paterson Institute) for the use of the fluorescence microscope, and Dr Steven Marsden (University of Manchester) for assistance with the AFM imaging. We also thank STFC Daresbury Laboratory, in particular James Nicholson, for the use of the Perkin Elmer Spotlight FTIR microscope.

References
